



МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ  
Федеральное государственное автономное образовательное учреждение высшего образования  
**«Дальневосточный федеральный университет»**  
(ДФУ)  
ИНСТИТУТ МАТЕМАТИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ

СОГЛАСОВАНО  
Руководитель ОП

Бедрина С.Л.

УТВЕРЖДАЮ

Заведующий департаментом Информационных и компьютерных систем

Пустовалов Е.В.

« 10 » марта

2022г.



**РАБОЧАЯ ПРОГРАММА ДИСЦИПЛИНЫ**  
Теоретические основы и технология обработки больших данных  
**Направление подготовки 09.03.03 Прикладная информатика**  
(Прикладная информатика в экономике)  
**Форма подготовки очная**

курс 3,4 семестр 7,8  
лекции 34 час.  
практические занятия 00 час  
лабораторные работы 68 час.  
в том числе с использованием МАО лек.      - / пр.      - / лаб. 34 час.  
всего часов аудиторной нагрузки 102 час.  
в том числе с использованием МАО 34 час.  
самостоятельная работа 258 час.  
в том числе на подготовку к экзамену 27 час.  
контрольные работы (количество) не предусмотрены  
курсовая работа / курсовой проект не предусмотрены  
зачет 6 семестр  
экзамен 7 семестр

Рабочая программа составлена в соответствии с требованиями Федерального государственного образовательного стандарта по направлению подготовки 09.03.03 Прикладная информатика утвержденного приказом Министерства образования и науки РФ от 19 сентября 2017 г. № 922

Рабочая учебная программа обсуждена на заседании департамента информационных и компьютерных систем протокол № 7 от «25» февраля 2022 г.

Заведующий департаментом Информационных и компьютерных систем: д.ф.-м. н., доцент Пустовалов Е.В.  
Составитель ст.преподаватель доцент, канд. техн. наук Оськин Д.А., Е.И.Шувалова

Владивосток  
2022

**Оборотная сторона титульного листа РПД**

**I. Рабочая программа пересмотрена на заседании кафедры/департамента:**

Протокол от « \_\_\_\_\_ » \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Заведующий кафедрой \_\_\_\_\_  
(подпись) (И.О. Фамилия)

**II. Рабочая программа пересмотрена на заседании кафедры/департамента:**

Протокол от « \_\_\_\_\_ » \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Заведующий кафедрой \_\_\_\_\_  
(подпись) (И.О. Фамилия)

**III. Рабочая программа пересмотрена на заседании кафедры/департамента:**

Протокол от « \_\_\_\_\_ » \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Заведующий кафедрой \_\_\_\_\_  
(подпись) (И.О. Фамилия)

**IV. Рабочая программа пересмотрена на заседании кафедры/департамента:**

Протокол от « \_\_\_\_\_ » \_\_\_\_\_ 20\_\_ г. № \_\_\_\_\_

Заведующий кафедрой \_\_\_\_\_  
(подпись) (И.О. Фамилия)

## **Цели и задачи освоения дисциплины:**

Цель: изучение некоторых современных информационных технологий, предназначенных для интеллектуального анализа данных, направленных на формирование целостного представления об анализе и интерпретации экспериментальных и статистических данных, как о процессе поиска, так и применения скрытых в них закономерностей для достижения поставленных целей.

Задачи:

- получить представления о феномене больших данных, о научных и технических проблемах и возможностях, связанных с их появлением, о трендах в области технологий хранения и анализа больших данных;
- приобрести знания причин возникновения тренда больших данных, процессов анализа больших данных, основных подходов к обработке больших массивов данных;
- приобрести умения формулировать алгоритмы, выбрать подходящий инструмент анализа больших данных, выбрать подходящую технологию хранения больших данных.

Для успешного изучения дисциплины «Теоретические основы и технология обработки больших данных» у обучающихся должны быть сформированы следующие предварительные компетенции:

- УК-1 Способен осуществлять поиск, критический анализ и синтез информации, применять системный подход для решения поставленных задач;
- УК-2 Способен определять круг задач в рамках поставленной цели и выбирать оптимальные способы их решения, исходя из действующих правовых норм, имеющихся ресурсов и ограничений;
- ОПК-2 Способен использовать современные информационные технологии и программные средства, в том числе отечественного производства, при решении задач профессиональной деятельности;
- ОПК-3 Способен решать стандартные задачи профессиональной деятельности на основе информационной и библиографической культуры с применением информационно-коммуникационных технологий и с учетом основных требований информационной безопасности;
- ОПК-5 Способен устанавливать программное и аппаратное обеспечение для информационных и автоматизированных систем.

В результате изучения данной дисциплины у обучающихся формируются следующие профессиональные компетенции:

Тип задач	Код и наименование профессиональной компетенции (результат освоения)	Код и наименование индикатора достижения компетенции
производственно-технологический	ПК-5 Способен интегрировать компоненты и сервисы информационных систем	ПК-5.1 Определяет современные методы интеграции компонентов и сервисов информационных систем ПК-5.2 Применяет наиболее эффективные решения интеграции для предприятия ПК-5.3 Применяет программные продукты интеграции компонентов и сервисов информационных систем

Код и наименование индикатора достижения компетенции	Наименование показателя оценивания (результата обучения по дисциплине)
ПК-5.1 Определяет современные методы интеграции компонентов и сервисов информационных систем	Знать: современные методы интеграции компонентов и сервисов информационных систем Уметь: определять современные методы интеграции компонентов и сервисов информационных систем Владеть: современными методами интеграции компонентов и сервисов информационных систем
ПК-5.2 Применяет наиболее эффективные решения интеграции для предприятия	Знать: методы интеграции для предприятия Уметь: применять наиболее эффективные решения интеграции для предприятия Владеть: методами и средствами определения наиболее эффективных решений интеграции для предприятия
ПК-5.3 Применяет программные продукты интеграции компонентов и сервисов информационных систем	Знать: программные продукты интеграции компонентов и сервисов информационных систем Уметь: применять программные продукты интеграции компонентов и сервисов информационных систем Владеть: программными продуктами интеграции компонентов и сервисов информационных систем

## 2. Трудоёмкость дисциплины и видов учебных занятий по дисциплине

Общая трудоёмкость дисциплины составляет 5 зачётных единиц, 180 академических часа).

(1 зачетная единица соответствует 36 академическим часам)

Видами учебных занятий и работы обучающегося по дисциплине являются:

Обозначение	Виды учебных занятий и работы обучающегося
Лек	Лекции
Лаб	Лабораторные работы
ОК	Онлайн-курс
СР	Самостоятельная работа обучающегося в период теоретического обучения
Контроль	Самостоятельная работа обучающегося и контактная работа обучающегося с

## Структура дисциплины:

Форма обучения – очная.

№	Наименование раздела дисциплины	Семестр	Количество часов по видам учебных занятий и работы обучающегося					Формы промежуточной аттестации	
			Лек	Лаб	Пр	ОК	СР		Контроль
1	<b>Введение в интеллектуальный анализ больших данных</b>	6	4						УО-1 ПР-6
2	<b>Мягкие вычисления в анализе больших данных</b>	6	12			72	56		
3	<b>Распределенный анализ данных</b>	6	2						
	Итого 6 семестр		18	34		72	56		
4	<b>Классические подходы Data Mining</b>	7	16				103	27	
	Итого 7 семестр		16	34			103		
	Итого:		34	68		72	159	27	

**I. СТРУКТУРА И СОДЕРЖАНИЕ ТЕОРЕТИЧЕСКОЙ ЧАСТИ КУРСА****Лекционные занятия (34 час.)***6 Семестр (18часов)***Раздел I. Введение в интеллектуальный анализ больших данных ( 4 час.)****Тема 1. Большие данные. Хранилища данных ( 2 час.)**

Большие данные и проблемы их обработки. Методы представления больших данных. Концепция хранилищ данных. Организация Хранилищ данных. Многомерная модель данных. Определение OLAP-системы. Архитектура OLAP-систем: MOLAP, ROLAP, HOLAP.

**Тема 2. Введение в интеллектуальный анализ (Data Mining) ( 2 час.)**

Анализ данных методами Data Mining. Задачи Data Mining: классификация, регрессионный анализ, кластеризация, поиск ассоциативных правил. Области применения Data Mining. Модели Data Mining: предсказательные и описательные. Технологии и методы обработки больших данных: классические

(классификация, регрессия, поиск ассоциативных правил, кластеризация) и интеллектуальные (нейронные сети, нечеткие или размытые (fuzzy) алгоритмы, генетические алгоритмы) подходы. Обнаружение знаний.

## **Раздел II. Мягкие вычисления в анализе больших данных ( 12 час.)**

### **Тема 1. Нейросетевые технологии анализа больших данных ( 4 час.)**

#### **Нейронные сети. Базовые понятия.**

Биологические представления о нейроне. Понятие нейрокомпьютера. Классификация нейронных сетей. Задача распознавания и линейная машина. Искусственный нейрон. Проблема линейной разделимости. Правило обучения Хебба. Концепция входной и выходной звезды. Парадигмы обучения. Предварительная обработка информации и оценка качества работы нейросети.

#### **Однослойные нейронные сети.**

Описание искусственного нейрона. Персептрон. Линейная нейронная сеть. Рекуррентный метод наименьших квадратов. Линейная сеть с линией задержки.

#### **Нейронные сети прямого распространения.**

Топология и свойства. Алгоритм обратного распространения ошибки. Реализация логических функций. Аппроксимация функций. Распознавание символов. Моделирование статических зависимостей. Масштабирование и восстановление данных.

#### **Модели ассоциативной памяти.**

Нейронная сеть Элмана. Сети Хопфилда. Двухнаправленная ассоциативная память. Нейронная сеть Хэмминга. Адаптивные резонансные нейронные сети.

#### **Нейронные сети Кохонена.**

Структура сети Кохонена. Обучение сети Кохонена. Слой Кохонена. Самоорганизующиеся карты Кохонена. Нейронные сети классификации.

#### **Радиальные нейронные сети.**

Структура радиальной нейронной сети. Расчет параметров радиальной нейронной сети. Обучение радиальной нейронной сети. Радиальные нейронные сети и нечеткие системы.

#### **Методы прогнозирования на базе нейронных сетей**

#### **Методы классификации и кластеризации на базе нейронных сетей**

### **Тема 2. Технологии размытой (нечеткой) логики в анализе больших данных ( 4 час.)**

Введение в теорию множеств. Операции над множествами.

Введение в теорию нечетких множеств. Операции над нечеткими множествами. Нечеткие отношения. Нечеткие отображения.

Нечеткость в описании прикладных задач. Нечеткость в принятии решений. Нечеткие множества в анализе данных.

### **Тема 3. Генетические алгоритмы ( 2 час.)**

Понятие о генетическом алгоритме. Этапы работы генетического алгоритма. Кодирование информации и формирование популяции. Оценивание популяции. Селекция. Скрещивание и формирование нового поколения. Мутация. Настройка параметров генетического алгоритма. Канонический генетический алгоритм. Пример работы генетического алгоритма. Рекомендации к программной реализации генетического алгоритма. Применение генетического алгоритма для решения задач оптимизации и аппроксимации.

### **Тема 4. Анализ текстовой информации ( 2 час.)**

Задачи анализа текстов: этапы анализа текстов, преобразование текстов. Извлечение ключевых понятий из текста: описание процесса извлечения понятий из текста; стадия локального анализа; стадия интеграции и вывода понятий.

Классификация текстовых документов: задачи и методы классификации текстов.

Методы кластеризации текстовых документов: представление текстовых документов, иерархические и бинарные методы кластеризации.

Задачи аннотирования текстов: выполнение аннотирования текстов; методы извлечения фрагментов для аннотации.

Средства анализа текстовой информации: Oracle Text (Oracle), Intelligent Miner for Text (IBM), Text Miner (SAS Institute), TextAnalyst (Мегacomпьютер Интеллидженс).

## **Раздел III. Распределенный анализ данных ( 2 час.)**

### **Тема 1. Мультиагентный подход к распределенному анализу данных ( 2 час.)**

Системы мобильных агентов: стандарты мультиагентных систем, системы мобильных агентов.

Использование мобильных агентов для анализа данных: проблемы распределенного анализа данных, агенты-аналитики, варианты анализа распределенных данных).

Система анализа распределенных данных: общий подход, агент для сбора информации о базе данных, агент для сбора статистической информации о

данных, агент для решения задачи интеллектуального анализа данных, агент для решения интегрированной задачи интеллектуального анализа данных.

*7 Семестр (18 часов)*

#### **Раздел IV. Классические подходы Data Mining ( 16 час.)**

##### **Тема 1 Введение в интеллектуальный анализ (Data Mining) ( 2 час.)**

Анализ данных методами Data Mining. Задачи Data Mining: классификация, регрессионный анализ, кластеризация, поиск ассоциативных правил. Области применения Data Mining. Модели Data Mining: предсказательные и описательные. Технологии и методы обработки больших данных: классические (классификация, регрессия, поиск ассоциативных правил, кластеризация) и интеллектуальные (нейронные сети, нечеткие или размытые (fuzzy) алгоритмы, генетические алгоритмы) подходы. Обнаружение знаний.

**Deductor . Назначение. Общие сведения.**

##### **Тема 2 Задача классификации ( 2 час.)**

Постановка задачи. Представление результатов: правила классификации; деревья решений; математические функции.

Методы построения правил классификации: построение элементарных правил; методика Байеса.

Методы построения деревьев решений.

##### **Тема 3 Регрессионный анализ ( 2 час.)**

Общий вид математических зависимостей. Простая и множественная линейная регрессия. Метод наименьших квадратов. Ограничения применимости регрессионных моделей.

Нелинейные модели.

Прогнозирование временных рядов: постановка задачи, методы прогнозирования временных рядов.

##### **Тема 4 Поиск ассоциативных правил ( 4 час.)**

Постановка задачи, поиск последовательностей, обзор задач поиска ассоциативных правил. Представление результатов. Алгоритм Apriori: свойство антимонотонности, описание алгоритма.

##### **Тема 5 Кластеризация ( 4 час.)**

Постановка задачи. Меры близости. Представление результатов. Алгоритмы кластеризации: базовые (иерархические (агломеративные и



дивизимные), неиерархические (k-means, fuzzy c-means, Гюстафсона-Кесселя)) и адаптивные.

Модельные (алгоритм EM), концептуальные (алгоритм Cobweb), сетевые (метод WaveCluster) алгоритмы.

### **Тема 6 Стандарты Data Mining ( 2 час.)**

Обзор стандартов. Стандарт CWM: назначение, структура и состав. Стандарт CRISP: структура, фазы и задачи стандарта. Стандарт PMML. Другие стандарты: SQL/MM, OLE DB, JDMAPI.

## **II. СТРУКТУРА И СОДЕРЖАНИЕ ПРАКТИЧЕСКОЙ ЧАСТИ КУРСА И САМОСТОЯТЕЛЬНОЙ РАБОТЫ**

### **Лабораторные работы (68часов)**

#### **Лабораторные работы ( 34 час.)**

*VI Семестр (18 часов)* Применяются МАО лабораторный метод, творческое задание и работа в малых группах(18 час.)

**Лабораторная работа №1.** Создание модели многослойной нейронной сети, оценка параметров многослойной нейронной сети прямого распространения градиентным алгоритмом обучения (для построения статической модели) (4 час.)

Цель работы: получение навыков по формированию обучающей выборки и оценке параметров многослойной нейронной сети прямого распространения градиентным алгоритмом обучения (для построения статической модели линейных и нелинейных объектов). Разработка программного обеспечения в среде Scilab (или другой).

**Лабораторная работа №2.** Создание модели рекуррентной нейронной сети (многослойной с обратными связями, сети Элмана), оценка параметров рекуррентной нейронной сети градиентным алгоритмом обучения (для построения динамической модели) (6 час.)

Цель работы: получение навыков по формированию обучающей выборки и оценке параметров рекуррентной нейронной сети градиентным алгоритмом обучения (для построения динамической модели линейных и нелинейных объектов). Разработка программного обеспечения в среде Scilab (или другой).

**Лабораторная работа №3.** Программные реализации моделей нейронных сетей (10 час.)

Цель работы: получение навыков по моделированию радиально базисных сетей, сетей Кохонена и их обучению (применительно к использованию решения задач кластеризации). Моделирование работы сетей Хэмминга и Хопфилда. Исследование стохастических методов обучения нейронных сетей. Разработка программного обеспечения в среде Scilab (или другой).

**Лабораторная работа №4. Нечеткие системы (6 час.)**

Цель работы: создание нечеткого описания системы, изучение операций над нечеткими множествами.

**Лабораторная работа №5. Генетический алгоритм (6 час.)**

Цель работы: моделирование работы генетических алгоритмов для решения задач оптимизации.

**Лабораторная работа №6. Моделирование систем анализа распределенных данных в среде AnyLogic (4 час.)**

Цель работы: моделирование систем анализа распределенных данных в среде AnyLogic.

*VII Семестр (16 час.)* Применяются МАО лабораторный метод, творческое задание и работа в малых группах(16 час.)

**Лабораторная работа №1 Знакомство с аналитической платформой Deductor. Начало работы с системой (2час.)**

Цель работы: получение навыков установки и работы с проектами.

**Лабораторная работа №2 Визуализация данных. (2час.)**

Цель работы: получение навыков работы с *Маст ером визуализации*

**Лабораторная работа №3 Обработка данных в таблицах. (2час.)**

Цель работы: получение навыков обработки исходных данных

**Лабораторная работа №4 Работа с таблицами - сортировка записей, редактирование полей и данных. (2час.)**

Цель работы: получение навыков сортировки, замены, фильтрации данных.

**Лабораторная работа №5 Прогнозирование с помощью построения пользовательских моделей (2час.)**

Цель работы: получение навыков построения пользовательских моделей и прогнозирования с их использованием.

**Лабораторная работа №6 Классификация с помощью деревьев решений (2 час.)**

Цель работы: получение навыков решения задач классификации с помощью деревьев решений и прогнозирования с их использованием.

**Лабораторная работа №7 Кластеризация с помощью самоорганизующейся карты Кохонена (2час.)**

Цель работы: получение навыков решения задач кластеризации с помощью самоорганизующейся карты Кохонена и построения прогноза.

### **Лабораторная работа №8 Поиск ассоциативных правил (2час.)**

Цель работы: получение навыков нахождения закономерности между связанными событиями.

### **Задания для самостоятельной работы**

*Требования:* перед каждой лабораторной работой обучающемуся необходимо изучить Методические указания по выполнению лабораторных работ по дисциплине «Теоретические основы и технология обработки больших данных».

#### **Самостоятельная работа №1.**

*Требования:*

Проработать контрольные вопросы на тему **Высокопроизводительные вычисления:**

- 1.Что такое Apache Hadoop?
- 2.В чем преимущества решений на базе Hadoop?
- 3.Что такое MapReduce?
- 4.Какими достоинствами и недостатками обладает MapReduce?
- 5.Какому основному принципу следует HDFS?
- 6.Какой размер блока по умолчанию в HDFS?
- 7.Какие функции выполняет NameNode в HDFS?
- 8.Какой узел отвечает за репликацию данных в Hadoop?
- 9.Какие компоненты содержит Slave узел в Hadoop?
- 10.Какие компоненты содержит Master узел в Hadoop?
- 11.Какие компоненты являются частями HDFS?
- 12.Какое API было добавлено в Hadoop v2.0?
- 13.Для чего используется автономный режим Hadoop?
- 14.Какой режим необходим для того, чтобы на локальной машине использовать Hadoop как кластер, состоящий из одного узла?

#### **Самостоятельная работа №2.**

*Требования:*

Проработать контрольные вопросы по теме Сложные методы аналитики:

- 1.Чем анализ больших данных отличается от традиционного анализа?
- 2.Какие основные типы Data Mining?
- 3.Какие категории Web Mining можно выделить?

4. В чем основная задача Web Content Mining?
5. В чем основные задачи интеллектуального анализа текстов?
6. Что изучает статистика?
7. К каким алгоритмам классификации относится метод ближайших соседей?
8. Что является целью кластеризации?
9. С помощью какого алгоритма можно найти ассоциативное правило?
10. Что подразумевается под определением "статистический вывод"?
11. Чем отличаются ошибки первого и второго рода?

### III. УЧЕБНО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ САМОСТОЯТЕЛЬНОЙ РАБОТЫ ОБУЧАЮЩИХСЯ

Учебно-методическое обеспечение самостоятельной работы обучающихся по дисциплине включает в себя:

- план-график выполнения самостоятельной работы по дисциплине, в том числе примерные нормы времени на выполнение по каждому заданию;
- требования к представлению и оформлению результатов самостоятельной работы;
- критерии оценки выполнения самостоятельной работы.

#### План-график выполнения самостоятельной работы по дисциплине

№ п/п	Дата/сроки выполнения	Вид самостоятельной работы	Примерные нормы времени на выполнение	Форма контроля
1	В течение семестра	Подготовка к лабораторным занятиям, изучение литературы	34 часов	Работа на лабораторных занятиях (ПР-6)
2	В течение семестра	Выполнение самостоятельной работы № 1	22 часа	УО-1 (собеседование/устный опрос)
Итого за 6 семестр			56 час	
3	В течение семестра	Подготовка к лабораторным занятиям, изучение литературы	34 часов	Работа на лабораторных занятиях (ПР-6)
4	В течение семестра	Выполнение самостоятельной работы № 2	69 час	УО-1 (собеседование/устный опрос)
Итого за 7 семестр			103 час.	
5	15-16 неделя 7-го семестра	Подготовка к экзамену	27 часов	экзамен
Итого:			186 час.	

## **Рекомендации по самостоятельной работе студентов**

*Планирование и организация времени, отведенного на выполнение заданий самостоятельной работы.*

Изучив график выполнения самостоятельных работ, следует правильно её организовать. Рекомендуется изучить структуру каждого задания, обратить внимание на график выполнения работ, отчетность по каждому заданию предоставляется в последнюю неделю согласно графику. Обратите внимание, что итоги самостоятельной работы влияют на окончательную оценку по итогам освоения учебной дисциплины.

*Работа с литературой.*

При выполнении ряда заданий требуется работать с литературой. Рекомендуется использовать различные возможности работы с литературой: фонды научной библиотеки ДВФУ (<http://www.dvfu.ru/library/>) и других ведущих вузов страны, а также доступных для использования научно-библиотечных систем.

В процессе выполнения самостоятельной работы, в том числе при написании отчета рекомендуется работать со следующими видами изданий:

а) Научные издания, предназначенные для научной работы и содержащие теоретические, экспериментальные сведения об исследованиях. Они могут публиковаться в форме: монографий, научных статей в журналах или в научных сборниках;

б) Учебная литература подразделяется на:

- учебные издания (учебники, учебные пособия, тексты лекций), в которых содержится наиболее полное системное изложение дисциплины или какого-то ее раздела;

- справочники, словари и энциклопедии – издания, содержащие краткие сведения научного или прикладного характера, не предназначенные для сплошного чтения. Их цель – возможность быстрого получения самых общих представлений о предмете.

Существуют два метода работы над источниками:

– сплошное чтение обязательно при изучении учебника, глав монографии или статьи, то есть того, что имеет учебное значение. Как правило, здесь требуется повторное чтение, для того чтобы понять написанное. Старайтесь при сплошном чтении не пропускать комментарии, сноски, справочные материалы, так как они предназначены для пояснений и помощи. Анализируйте рисунки (карты, диаграммы, графики), старайтесь понять, какие тенденции и закономерности они отражают;

– метод выборочного чтения дополняет сплошное чтение; он применяется для поисков дополнительных, уточняющих необходимых сведений в словарях,

энциклопедиях, иных справочных изданиях. Этот метод крайне важен для повторения изученного и его закрепления, особенно при подготовке к зачету.

Для того чтобы каждый метод принес наибольший эффект, необходимо фиксировать все важные моменты, связанные с интересующей Вас темой.

Тезисы – это основные положения научного труда, статьи или другого произведения, а возможно, и устного выступления; они несут в себе большой объем информации, нежели план. Простые тезисы лаконичны по форме; сложные – помимо главной авторской мысли содержат краткое ее обоснование и доказательства, придающие тезисам более весомый и убедительный характер. Тезисы прочитанного позволяют глубже раскрыть его содержание; обучаясь излагать суть прочитанного в тезисной форме, вы сумеете выделять из множества мыслей авторов самые главные и ценные и делать обобщения.

Конспект – это способ самостоятельно изложить содержание книги или статьи в логической последовательности. Конспектируя какой-либо источник, надо стремиться к тому, чтобы немногими словами сказать о многом. В тексте конспекта желательно поместить не только выводы или положения, но и их аргументированные доказательства (факты, цифры, цитаты).

Писать конспект можно и по мере изучения произведения, например, если прорабатывается монография или несколько журнальных статей.

Составляя тезисы или конспект, всегда делайте ссылки на страницы, с которых вы взяли конспектируемое положение или факт, – это поможет вам сократить время на поиск нужного места в книге, если возникает потребность глубже разобраться с излагаемым вопросом или что-то уточнить при написании письменных работ.

### **Методические рекомендации по выполнению заданий для самостоятельной работы и критерии оценки.**

*Работа на лабораторных занятиях.* Отчетность по теме осуществляется в форме отчета. Отчет, как оценочное средство, позволяет оценить умение обучающегося письменно излагать суть поставленного вопроса, самостоятельно проводить анализ, формулировать выводы. Отчет предоставляется в письменном виде. Методические рекомендации по написанию отчета представлены ниже.

#### Критерии оценки.

<b>Оценка</b>	<b>Требования</b>
<b>«зачтено»</b>	Студент владеет навыками самостоятельной работы по теме исследования, реферировать литературные источники; методами анализа теоретических и/или практических аспектов изучаемой области. Отчет характеризуется смысловой цельностью, связностью и

	последовательностью изложения. Студент умеет обобщать фактический материал, делать самостоятельные выводы. Работа соответствует требованиям и выполнена в установленные сроки.
<b>«не зачтено»</b>	Не раскрыта структура и теоретическая составляющая темы. Студент не умеет обобщать фактический материал, делать самостоятельные выводы, не владеет навыком реферировать литературные источники. Отчет не выполнен.

*Самостоятельная работа №1-2.* От обучающегося требуется:

1. Свободно ориентироваться в предметной области.
2. Знать и уметь объяснить основные понятия и положения в соответствии с контрольными вопросами.

Собеседование (устный опрос) позволяет оценить знания и кругозор студента, умение логически построить ответ, владение монологической речью и иные коммуникативные навыки.

Опрос – важнейшее средство развития мышления и речи. Обучающая функция опроса состоит в выявлении деталей, которые по каким-то причинам оказались недостаточно осмысленными в ходе учебных занятий и при подготовке задания по самостоятельной работе.

Критерии оценки. Используется зачетная система. Во время опроса допускается не более 1-й ошибки или неточности по названию периода, его времени и длительности.

### **Методические рекомендации к оформлению отчетов по лабораторным работам**

Отчеты представляются в электронной форме, оформленные в MSWord по правилам, принятым в ДВФУ. Отчет должен содержать:

1. Титульный лист\*.
2. Содержание\*.
3. Задание.
4. Краткое изложение теоретических положений, необходимых для выполнения работы.
5. Основная часть: материалы выполнения заданий.
6. Анализ полученных результатов.
7. Выводы по работе\* (какие задачи решены, что освоено при выполнении работы).
8. Приложения\* (при необходимости, обычно в Приложении выносят листинг программы)

\* Включаются в отчет с новой страницы

#### IV. КОНТРОЛЬ ДОСТИЖЕНИЯ ЦЕЛЕЙ КУРСА

№ п/п	Контролируемые модули/разделы / темы дисциплины	Код индикатора достижения компетенции	Результаты обучения	Оценочные средства – наименование	
				текущий контроль	промежуточная аттестация
1	Введение в интеллектуальный анализ больших данных	УО-1 ПР-6	<p>знает основные понятия, классификацию и характеристики больших данных</p> <p>умеет реализовывать модели данных</p> <p>владеет терминологией больших данных</p>	собеседование (УО-1)	ПР-1
2	Мягкие вычисления в анализе больших данных	УО-1 ПР-6	<p>Знает методы синтеза и анализа нейросетевых, нечетких систем и генетические алгоритмы</p> <p>умеет синтезировать и применять на практике нейросетевые, нечеткие системы анализа данных</p> <p>владеет математическим аппаратом теории систем нейронных сетей, нечеткой логики</p>	лабораторная работа (ПР-6)	
3	Распределенный анализ данных	УО-1 ПР-6	<p>Знает методы мультиагентного моделирования</p> <p>Умеет использовать программные продукты для анализа распределенных данных</p> <p>Владеет математическим аппаратом описания распределенных систем анализа данных</p>	лабораторная работа (ПР-6)	
4	Классические подходы Data Mining	УО-1 ПР-6	<p>знает назначение и классы интеллектуальных информационных систем; модели представления знаний;</p> <p>умеет разрабатывать модель знаний</p>	лабораторная работа (ПР-6)	



			прикладной области; выбирать инструментальные средства и технологии проектирования экспертных систем; создавать и сопровождать большие базы данных		
			владеет навыками работы с инструментальными средствами представления модели знаний предметной области		

Типовые контрольные задания, методические материалы, определяющие процедуры оценивания знаний, умений и навыков и (или) опыта деятельности, а также качественные критерии оценивания, которые описывают уровень сформированности компетенций, представлены в разделе VIII.

## V. СПИСОК УЧЕБНОЙ ЛИТЕРАТУРЫ И ИНФОРМАЦИОННО-МЕТОДИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ

### Основная литература

1. Нечеткие множества и нейронные сети : учебное пособие / Яхьяева Г.Э.. — Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2020. — 315 с. — ISBN 978-5-4497-0665-2. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/97552.html> (дата обращения: 13.04.2021). — Режим доступа: для авторизир. пользователей
2. Генетические алгоритмы : учебник / Л. А. Гладков, В. В. Курейчик, В. М. Курейчик ; под ред. В. М. Курейчика. — 2-е изд., исправл. и доп. — Москва : ФИЗМАТЛИТ, 2010. - 368 с. - ISBN 978-5-9221-0510-1. - Текст : электронный. - URL: <https://znanium.com/catalog/product/544626> (дата обращения: 13.04.2021). — Режим доступа: по подписке.

### Дополнительная литература

1. Маккинли Уэс Python и анализ данных / Маккинли Уэс. — Саратов : Профобразование, 2019. — 482 с. — ISBN 978-5-4488-0046-7. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/88752.html> (дата обращения: 13.04.2021). — Режим доступа: для авторизир. пользователей

2. Проектирование систем искусственного интеллекта : учебное пособие / Сотник С.Л. — Москва : Интернет-Университет Информационных Технологий (ИНТУИТ), Ай Пи Ар Медиа, 2021. — 228 с. — ISBN 978-5-4497-0868-7. — Текст : электронный // Электронно-библиотечная система IPR BOOKS : [сайт]. — URL: <http://www.iprbookshop.ru/102054.html> (дата обращения: 13.04.2021). — Режим доступа: для авторизир. пользователей

#### **Перечень ресурсов информационно-телекоммуникационной сети «Интернет»**

1. Интернет-библиотека образовательных изданий: <http://www.iqlib.ru>
2. Интернет университет информационных технологий:  
<http://www.intuit.ru/>
3. Научная электронная библиотека eLIBRARY.RU:  
<https://elibrary.ru/defaultx.asp>
4. Информационная система «Единое окно доступа к образовательным ресурсам»: <http://window.edu.ru/window/library>
5. Электронно-библиотечная система Znanium.com (ООО "Знаниум"):  
<http://znanium.com/>
6. Электронно-библиотечная система ЮРАЙТ  
<https://www.biblio-online.ru/>
7. Электронно-библиотечная система «Консультант студента»  
<http://www.studentlibrary.ru/>
8. Электронно-библиотечная система «РУКОНТ» <https://lib.rucont.ru/>
9. Электронно-библиотечная система «IPRBOOKS»  
<http://www.iprbookshop.ru/>

#### **Перечень информационных технологий и программного обеспечения**

1. Сайт системы моделирования python: <https://www.python.org/downloads/>

#### **Профессиональные базы данных и информационные справочные системы**

1. База данных Scopus <http://www.scopus.com/home.url>
2. База данных Web of Science <http://apps.webofknowledge.com/>
3. Российский индекс национального цитирования  
<https://www.elibrary.ru/defaultx.asp>

## **VI. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

### **Планирование и организация времени, отведенного на изучение**

**дисциплины.** Приступить к освоению дисциплины следует незамедлительно в самом начале учебного семестра. Рекомендуется изучить структуру и основные положения Рабочей программы дисциплины. Обратит внимание, что кроме аудиторной работы (лекции, лабораторные занятия) планируется самостоятельная работа, итоги которой влияют на окончательную оценку по итогам освоения учебной дисциплины. Все задания (аудиторные и самостоятельные) необходимо выполнять и предоставлять на оценку в соответствии с графиком.

В процессе изучения материалов учебного курса предлагаются следующие формы работ: чтение лекций, лабораторные занятия, задания для самостоятельной работы.

*Лекционные занятия* ориентированы на освещение вводных тем в каждый раздел курса и призваны ориентировать студентов в предлагаемом материале, заложить научные и методологические основы для дальнейшей самостоятельной работы студентов.

*Лабораторные занятия* акцентированы на наиболее принципиальных и проблемных вопросах курса и призваны стимулировать выработку практических умений.

Особо значимой для профессиональной подготовки студентов является *самостоятельная работа* по курсу. В ходе этой работы студенты отбирают необходимый материал по изучаемому вопросу и анализируют его. Студентам необходимо ознакомиться с основными источниками, без которых невозможно полноценное понимание проблематики курса.

Освоение курса способствует развитию навыков обоснованных и самостоятельных оценок фактов и концепций. Поэтому во всех формах контроля знаний, особенно при сдаче зачета, внимание обращается на понимание проблематики курса, на умение практически применять знания и делать выводы.

**Работа с литературой.** Рекомендуется использовать различные возможности работы с литературой: фонды научной библиотеки ДВФУ и электронные библиотеки (<http://www.dvfu.ru/library/>), а также доступные для использования другие научно-библиотечные системы.

**Подготовка к экзамену.** К сдаче экзамена допускаются обучающиеся, выполнившие все задания (лабораторные, самостоятельные), предусмотренные учебной программой дисциплины, посетившие не менее 85% аудиторных занятий.

## **VII. МАТЕРИАЛЬНО-ТЕХНИЧЕСКОЕ ОБЕСПЕЧЕНИЕ ДИСЦИПЛИНЫ**

Перечень материально-технического и программного обеспечения дисциплины приведен в таблице.

### Материально-техническое и программное обеспечение дисциплины

Наименование специальных помещений и помещений для самостоятельной работы	Оснащенность специальных помещений и помещений для самостоятельной работы	Перечень лицензионного программного обеспечения. Реквизиты подтверждающего документа
690922, Приморский край, г. Владивосток, остров Русский, полуостров Саперный, поселок Аякс, 10, корпус D, ауд. D 533. Учебная аудитория для проведения занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации	Помещение укомплектовано специализированной учебной мебелью (посадочных мест – 20) Оборудование: ЖК-панель 47", FullHD, LGM4716 CCBA – 1 шт., Персональный компьютер 20 Доска аудиторная, Проектор	Пакет прикладных программ Python Пакет прикладных программ AnyLogic Пакет прикладных программ Deductor
690922, Приморский край, г. Владивосток, остров Русский, полуостров Саперный, поселок Аякс, 10, корпус D, ауд. D 534. Учебная аудитория для проведения занятий лекционного типа, групповых и индивидуальных консультаций, текущего контроля и промежуточной аттестации	Помещение укомплектовано специализированной учебной мебелью (посадочных мест – 20) Оборудование: ЖК-панель 47", FullHD, LGM4716 CCBA – 1 шт., Персональный компьютер 20 Доска аудиторная, Проектор	Пакет прикладных программ Python Пакет прикладных программ AnyLogic Пакет прикладных программ Deductor

Для проведения учебных занятий по дисциплине, а также для организации самостоятельной работы студентам доступно следующее лабораторное оборудование и специализированные кабинеты, соответствующие действующим санитарным и противопожарным нормам, а также требованиям техники безопасности при проведении учебных и научно-производственных работ.

В целях обеспечения специальных условий обучения инвалидов и лиц с ограниченными возможностями здоровья в ДВФУ все здания оборудованы пандусами, лифтами, подъемниками, специализированными местами, оснащенными туалетными комнатами, табличками информационно-навигационной поддержки.

## VIII. ФОНДЫ ОЦЕНОЧНЫХ СРЕДСТВ

Для дисциплины «Теоретические основы и технология обработки больших данных» используются следующие оценочные средства:

Устный опрос:

1. Собеседование (УО-1)

Письменные работы:

1. Лабораторная работа (ПР-6)

### **Устный опрос**

Устный опрос позволяет оценить знания и кругозор студента, умение логически построить ответ, владение монологической речью и иные коммуникативные навыки.

Обучающая функция состоит в выявлении деталей, которые по каким-то причинам оказались недостаточно осмысленными в ходе учебных занятий и при подготовке к зачёту.

Собеседование (УО-1) – средство контроля, организованное как специальная беседа преподавателя с обучающимся на темы, связанные с изучаемой дисциплиной, и рассчитанное на выяснение объема знаний обучающегося по определенному разделу, теме, проблеме и т.п.

### **Письменные работы**

Письменный ответ приучает к точности, лаконичности, связности изложения мысли. Письменная проверка используется во всех видах контроля и осуществляется как в аудиторной, так и во внеаудиторной работе.

Лабораторная работа (ПР-6) – средство для закрепления и практического освоения материала по определенному разделу.

## **Методические рекомендации, определяющие процедуры оценивания результатов освоения дисциплины**

### **Оценочные средства для промежуточной аттестации**

Промежуточная аттестация студентов по дисциплине «Теоретические основы и технология обработки больших данных» проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной. Форма отчётности по дисциплине – зачет (6-й, весенний семестр) и экзамен (7-й, осенний семестр). Зачет по дисциплине включает ответы на 2 вопроса. Один из вопросов носит общий характер. Он направлен на раскрытие студентом знаний по «сквозным» вопросам сетевых технологий. Экзамен по дисциплине включает 3 вопроса.

### **Методические указания по сдаче зачета**

Зачет принимается ведущим преподавателем. При большом количестве групп у одного преподавателя или при большой численности потока по

распоряжению заведующего кафедрой (заместителя директора по учебной и воспитательной работе) допускается привлечение в помощь ведущему преподавателю других преподавателей. В первую очередь привлекаются преподаватели, которые проводили лабораторные занятия по дисциплине в группах.

В исключительных случаях, по согласованию с заместителем директора Школы по учебной и воспитательной работе, заведующий кафедрой имеет право принять зачет в отсутствие ведущего преподавателя.

Форма проведения зачета (устная, письменная и др.) утверждается на заседании кафедры по согласованию с руководителем в соответствии с рабочей программой дисциплины.

Во время проведения зачета студенты могут пользоваться рабочей программой дисциплины, а также с разрешения преподавателя, проводящего зачет, справочной литературой и другими пособиями (учебниками, учебными пособиями, рекомендованной литературой и т.п.).

Время, предоставляемое студенту на подготовку к ответу на зачете, должно составлять не более 20 минут. По истечении данного времени студент должен быть готов к ответу.

Присутствие на зачете посторонних лиц (кроме лиц, осуществляющих проверку) без разрешения соответствующих лиц (ректора либо проректора по учебной и воспитательной работе, директора Школы, руководителя ОПОП или заведующего кафедрой), не допускается. Инвалиды и лица с ограниченными возможностями здоровья, не имеющие возможности самостоятельного передвижения, допускаются зачет с сопровождающими.

При промежуточной аттестации обучающимся устанавливается оценка «зачтено» или «не зачтено».

В зачетную книжку студента вносится только запись «зачтено», запись «не зачтено» вносится только в экзаменационную ведомость. При неявке студента на зачет в ведомости делается запись «не явился».

### **Вопросы к зачету**

1. Что такое искусственная нейронная сеть? Что такое нейрокомпьютер и нейрокомпьютинг?
2. Какие принципы используются при классификации нейронных сетей?
3. Каковы три основные парадигмы обучения нейронных сетей? Чем отличается детерминированное обучение от стохастического?
4. Какие операции могут выполняться при предварительной обработке обучающих данных для нейросети?

5. Как оценить качество обучения нейросети?
6. Как используется МНК при обучении линейной нейронной сети? Как оценить качество обучения линейной ИНС при использовании МНК? Возможно ли использование МНК при обучении персептрона?
7. Как можно использовать нейронные сети ADALIN? Какие существуют ограничения при использовании ADALIN? Что представляют собой нейронные сети MADALINE?
8. Каковы свойства искусственных нейронных сетей прямого распространения? Каким свойством должна обладать ИНС прямого распространения, чтобы быть универсальным аппроксиматором?
9. Какова последовательность действий при использовании алгоритма обратного распространения ошибки? В чем состоит явление переобучения НС?
10. Что собой представляет градиентный метод поиска минимума функции?
11. В чем заключается основной недостаток алгоритма обратного распространения ошибки для обучения нейронных сетей?
12. Какова структура нейронной сети Элмана? Какие активационные функции используются в разных слоях сети Элмана?
13. Какую активационную функцию используют нейроны сети Хопфилда? Какова структура нейронной сети Хэмминга? Как вычисляется расстояние Хэмминга? Какие активационные функции используются в однослойной сети Хэмминга?
14. Какие задачи решают нейронные сети Кохонена? Сколько слоев имеет нейронная сеть Кохонена? Какие существуют варианты сетей Кохонена? Что такое топология карты Кохонена? Какие бывают варианты топологий? Каков механизм реализации обучения без учителя в сетях Кохонена? Какие проблемы могут возникать при обучении сетей Кохонена?
15. Какова основная идея использования генетического алгоритма для глобальной оптимизации? Что представляет собой хромосома при генетическом синтезе нейронной сети? Какие существуют генетические операторы? Что такое пригодность хромосомы и как ее можно оценить? В чем измеряется время работы генетического алгоритма?
16. Сколько слоев содержит RBF-сеть? Что представляет собой радиально-базисная функция? Является ли RBF-сеть универсальным аппроксиматором? Каковы возможные варианты радиально-базисных функций?
17. В чем состоит главное отличие RBF-сетей от обычных многослойных сетей прямого распространения?

18. Какую активационную функцию имеют нейроны выходного слоя RBF-сети?
19. Какой метод используется для приближенного выбора коэффициентов RBF-сети?
20. Какие параметры RBF-сети полагаются заданными при использовании метода наименьших квадратов?
21. С какой целью используется кластеризация при обучении RBF-сети?
22. Какие задачи можно решить с помощью генетического алгоритма при обучении RBF-сети?
23. Каковы принципы использования генетического алгоритма при обучении RBF-сети? Что представляют собой сети GRNN? Что представляют собой сети PNN?
24. Возникновение нечеткой логики. Нечеткие множества. Сущность и определения.
25. Дайте определение понятий «лингвистическая переменная» и «нечеткое множество», поясните их на примере. Какие операции можно выполнять над нечеткими множествами? Рассмотрите математические операции над нечеткими множествами: дополнение, объединение и пересечение множеств.
26. Операции над нечеткими множествами. Логические операции.
27. Операции над нечеткими множествами. Алгебраические операции. Построение функций принадлежности. Нечеткие и лингвистические переменные.
28. Дайте определение нечеткого отношения и расскажите о свойствах нечетких отношений. Каким образом, и с какой целью используются нечеткие отношения в ИИС?
29. Что такое нечеткая импликация? Какими способами ее можно реализовать для правил с одним выходом и двумя выходами. Приведите примеры.
30. Расскажите о нечетком логическом выводе. Дайте определение композиции нечетких отношений. Приведите пример нечеткой композиции.
31. Нечеткие алгоритмы и выводы. Расскажите об организации ЭС с нечетким логическим выводом. Какие способы используются в таких системах для представления и обработки знаний?
32. Программные пакеты в области нечеткой логики.
33. Среда моделирования Anylogic. Основные инструменты создания и работы с имитационной моделью.
34. Агентные модели. Особенности создания. Дать определение агента и описать его свойства.



35. Что называется многоагентной системой? Какие динамические свойства системы обнаруживаются в результате взаимодействия агентов?
36. Какие элементы динамической модели предприятия можно описать с помощью агентного моделирования? Области приложения агентного моделирования.

### **Критерии выставления оценки студенту на зачете**

К зачету допускаются обучающиеся, выполнившие программу обучения по дисциплине, прошедшие все этапы текущей аттестации.

<b>Оценка</b>	<b>Требования к сформированным компетенциям</b>
зачтено	<p>Оценка «зачтено» выставляется студенту при:</p> <ul style="list-style-type: none"> <li>- глубоко и прочном усвоении материала курса «Теоретические основы и технология обработки больших данных»,</li> <li>- исчерпывающе, последовательно, четко и логически стройно излагает материал,</li> <li>- умеет увязывать теорию с практикой,</li> <li>- использует в ответе материал монографической литературы,</li> <li>- правильно обосновывает принятое решение.</li> </ul>
не зачтено	<p>Оценка «неудовлетворительно» выставляется студенту:</p> <ul style="list-style-type: none"> <li>- который не знает значительной части материала,</li> <li>- допускает существенные ошибки,</li> <li>- неуверенно, с большими затруднениями выполняет лабораторные работы.</li> </ul>

### **Методические указания по сдаче экзамена**

Экзамен принимается ведущим преподавателем. При большом количестве групп у одного преподавателя или при большой численности потока по распоряжению заведующего кафедрой (заместителя директора по учебной и воспитательной работе) допускается привлечение в помощь ведущему преподавателю других преподавателей. В первую очередь привлекаются преподаватели, которые проводили лабораторные занятия по дисциплине в группах.

В исключительных случаях, по согласованию с заместителем директора Школы по учебной и воспитательной работе, заведующий кафедрой имеет право принять зачет в отсутствие ведущего преподавателя.

Форма проведения экзамена (устная, письменная и др.) утверждается на заседании кафедры по согласованию с руководителем в соответствии с рабочей программой дисциплины.

Во время проведения экзамена студенты могут пользоваться рабочей программой дисциплины, а также с разрешения преподавателя, проводящего зачет, справочной литературой и другими пособиями (учебниками, учебными пособиями, рекомендованной литературой и т.п.).

Время, предоставляемое студенту на подготовку к ответу на экзамене, должно составлять не более 60 минут. По истечении данного времени студент должен быть готов к ответу.

Присутствие на экзамене посторонних лиц (кроме лиц, осуществляющих проверку) без разрешения соответствующих лиц (ректора либо проректора по учебной и воспитательной работе, директора Школы, руководителя ОПОП или заведующего кафедрой), не допускается. Инвалиды и лица с ограниченными возможностями здоровья, не имеющие возможности самостоятельного передвижения, допускаются зачет с сопровождающими.

При промежуточной аттестации обучающимся устанавливается оценка «отлично», «хорошо», «удовлетворительно» или «не удовлетворительно».

В зачетную книжку студента вносится положительная оценка «отлично», «хорошо», «удовлетворительно», запись «не удовлетворительно» вносится только в экзаменационную ведомость. При неявке студента на экзамен в ведомости делается запись «не явился».

### **Вопросы к экзамену**

- 1 Определите понятие нечеткой логики.
- 2 Охарактеризуйте функцию принадлежности.
- 3 Опишите понятие дефазификации нечеткого множества.
- 4 Каким образом коэффициент уверенности выражается через меры доверия и недоверия?
- 5 Приведите соотношение между мерами доверия, полученными при независимом учете первого и второго свидетельства и объединенной мерой доверия, полученной при учете двух свидетельств.
- 6 Охарактеризуйте нечеткие правила вывода в экспертных системах.
- 7 Опишите структуру генетического алгоритма.
- 8 Охарактеризуйте целочисленное и вещественное кодирование в ГА.
- 9 Опишите канонический генетический алгоритм.
- 10 Охарактеризуйте операторы кроссовера и мутации.
- 11 Опишите биологические нейронные сети.
- 12 Охарактеризуйте понятие формального нейрона.
- 13 Опишите существующие модели искусственных нейронных сетей.
- 14 Приведите примеры активационных функций.
- 15 Опишите процесс обучения нейронной сети.
- 16 Интеллектуальный анализ данных (data mining).
- 17 Типы выявляемых закономерностей data mining: ассоциация, последовательность, классификация, кластеризация,

прогнозирование.

18 Математический аппарат data mining: деревья решений, регрессионный анализ, нейронные сети, временные ряды.

19 Области применения технологий интеллектуального анализа данных: розничная торговля, банковская деятельность, страховой бизнес, производство, автоматизированные системы для интеллектуального анализа данных.

### Критерии выставления оценки студенту на экзамене

К экзамену допускаются обучающиеся, выполнившие программу обучения по дисциплине, прошедшие все этапы текущей аттестации.

Оценка	Требования к сформированным компетенциям
«отлично»	Студент показал развернутый ответ, представляющий собой связное, логическое, последовательное раскрытие поставленного вопроса, широкое знание литературы. Студент показал понимание материала, обоснованность суждений, способность применить полученные знания на практике. Допускаются некоторые неточности в ответе, которые студент исправляет самостоятельно.
«хорошо»	Студент показал развернутый ответ, представляющий собой связное, логическое, последовательное раскрытие поставленного вопроса, есть неточности в ответе, которые студент не может исправить самостоятельно.
«удовлетворительно»	Студент показал понимание материала, обоснованность суждений, способность применить полученные знания на практике. Не может дать развернутого ответа. Есть неточности в ответе, которые студент не может исправить самостоятельно.
«не удовлетворительно»	Преподаватель обнаруживает незнание большей части проблем, связанных с изучением вопроса, допускает ошибки в ответе, искажает смысл текста, беспорядочно и неуверенно излагает материал. Данная оценка характеризует недостатки в подготовке студента, которые являются серьезным препятствием к успешной профессиональной и научной деятельности.

### Оценочные средства для текущей аттестации

Текущая аттестация студентов по дисциплине проводится в соответствии с локальными нормативными актами ДВФУ и является обязательной.

Текущая аттестация проводится в форме контрольных мероприятий (собеседования, презентации, эссе, лабораторных работ, контрольно-расчетных работ, творческого задания) по оцениванию фактических результатов обучения студентов и осуществляется ведущим преподавателем.

Объектами оценивания выступают:

– учебная дисциплина (активность на занятиях, своевременность

выполнения различных видов заданий, посещаемость всех видов занятий по аттестуемой дисциплине);

- степень усвоения теоретических знаний;
- уровень овладения практическими умениями и навыками по всем видам учебной работы;
- результаты самостоятельной работы.

Составляется календарный план контрольных мероприятий по дисциплине. Оценка посещаемости, активности обучающихся на занятиях, своевременность выполнения различных видов заданий ведётся на основе журнала, который ведёт преподаватель в течение учебного семестра.

### **Вопросы для собеседования / устного опроса**

#### **Раздел 1.**

- 1 Что означает термин «Big Data» в информационных технологиях?
- 2 Что является основной целью обработки Big Data?
- 3 Какие главные характеристики Big Data? Какие понятия содержит в себе принцип "V"?
- 4 Чем характеризуются "Большие данные"?
- 5 Какую модель данных использует большинство OLAP-систем?
- 6 Чем обусловлена неприменимость OLTP-систем для оперативного анализа данных?
- 7 Дайте определение ХД. Перечислите виды и задачи ХД.
- 8 Что такое витрины данных? Какова их роль в ХД?
- 9 Что такое таблицы фактов и измерений? Поясните понятия «схема звезда» и «схема снежинка».
- 10 Опишите многомерную модель данных.
- 11 Дайте определение OLAP. Опишите задачи OLAP.
- 12 Перечислите правила Кодда для OLAP-систем.
- 13 Что такое тест FASMI?
- 14 Приведите пример архитектур OLAP-систем.
- 15 Что такое ROLAP, MOLAP, HOLAP, DOLAP?
- 16 Дайте определение Data Mining. Приведите классификацию задач Data Mining.
- 17 Каково соотношение Data Mining и OLAP?

#### **Раздел 2.**

- 18 Определите понятие нечеткой логики.
- 19 Охарактеризуйте функцию принадлежности.

- 21 Опишите понятие дефазификации нечеткого множества.
- 22 Каким образом коэффициент уверенности выражается через меры доверия и недоверия?
- 23 Приведите соотношение между мерами доверия, полученными при независимом учете первого и второго свидетельства и объединенной мерой доверия, полученной при учете двух свидетельств.
- 24 Охарактеризуйте нечеткие правила вывода в экспертных системах.
- 25 Опишите структуру генетического алгоритма.
- 26 Охарактеризуйте целочисленное и вещественное кодирование в ГА.
- 27 Опишите канонический генетический алгоритм.
- 28 Охарактеризуйте операторы кроссовера и мутации.
- 29 Опишите биологические нейронные сети.
- 30 Охарактеризуйте понятие формального нейрона.
- 31 Опишите существующие модели искусственных нейронных сетей.
- 32 Приведите примеры активационных функций.
- 33 Опишите процесс обучения нейронной сети.

### **Раздел 3.**

- 34 Среда моделирования Anylogic. Основные инструменты создания и работы с имитационной моделью.
- 35 Агентные модели. Особенности создания. Дать определение агента и описать его свойства.
- 36 Что называется многоагентной системой? Какие динамические свойства системы обнаруживаются в результате взаимодействия агентов?
- 37 Какие элементы динамической модели предприятия можно описать с помощью агентного моделирования? Области приложения агентного моделирования.

### **Раздел 4.**

- 38 Интеллектуальный анализ данных (data mining).
- 39 Типы выявляемых закономерностей data mining: ассоциация, последовательность, классификация, кластеризация, прогнозирование.
- 40 Математический аппарат data mining: деревья решений, регрессионный анализ, нейронные сети, временные ряды.
- 41 Области применения технологий интеллектуального анализа данных: розничная торговля, банковская деятельность, страховой бизнес, производство, автоматизированные системы для интеллектуального анализа данных.

## Критерии оценивания

Оценка	Требования
<b>«зачтено»</b>	Студент показал развернутый ответ на вопрос, знание литературы, обнаружил понимание материала, обоснованность суждений, неточности в ответе исправляет самостоятельно.
<b>«не зачтено»</b>	Студент обнаруживает незнание вопроса, неуверенно излагает ответ.

## Тематика лабораторных работ

1. Создание модели многослойной нейронной сети, оценка параметров многослойной нейронной сети прямого распространения градиентным алгоритмом обучения (для построения статической модели)
2. Создание модели рекуррентной нейронной сети (многослойной с обратными связями, сети Элмана), оценка параметров рекуррентной нейронной сети градиентным алгоритмом обучения (для построения динамической модели)
3. Программные реализации моделей нейронных сетей
4. Нечеткие системы
5. Генетический алгоритм
6. Моделирование систем анализа распределенных данных в среде AnyLogic
7. Знакомство с аналитической платформой Deductor. Начало работы с системой
8. Визуализация данных.
9. Обработка данных в таблицах.
10. Работа с таблицами - сортировка записей, редактирование полей и данных.
11. Прогнозирование с помощью построения пользовательских моделей
12. Классификация с помощью деревьев решений
13. Кластеризация с помощью самоорганизующейся карты Кохонена
14. Поиск ассоциативных правил

## Критерии оценки лабораторных работ

Оценка	Требования
<b>«зачтено»</b>	Студент выполняет лабораторную работу в полном объеме с соблюдением необходимой последовательности проведения измерений, правильно самостоятельно определяет цель работы; самостоятельно, рационально выбирает необходимое оборудование для получения наиболее точных результатов проводимой работы. Грамотно и логично описывает ход работы, правильно формулирует выводы, точно и аккуратно выполняет все записи, таблицы, рисунки,

	чертежи, графики, вычисления и т.п., умеет обобщать фактический материал. Допускается два/три недочёта или одна негрубая ошибка и один недочёт. Работа соответствует требованиям и выполнена в срок.
<i>«не зачтено»</i>	Студент выполнил работу не полностью, объём выполненной части не позволяет сделать правильные выводы; не определяет самостоятельно цель работы; в ходе работы допускает одну и более грубые ошибки, которые не может исправить, или неверно производит наблюдения, измерения, вычисления и т.п.; не умеет обобщать фактический материал. Лабораторная работа не выполнена.